

Computationally efficient spatial modeling of annual maximum 24 hour precipitation. An application to data from Iceland.

Óli Páll Geirsson and Birgir Hrafnkelsson

Department of Mathematics

Faculty of Physical Sciences

School of Engineering and Natural Sciences

University of Iceland,

Iceland

and

Daniel Simpson

Department of Mathematical Sciences

Norwegian University of Science and Technology

Norway

May 28, 2014

Abstract

We propose a computationally efficient statistical method to obtain distributional properties of annual maximum 24 hour precipitation on a 1 km by 1 km regular grid over Iceland. A latent Gaussian model is built which takes into account observations, spatial variations and outputs from a local meteorological model. A covariate based on the meteorological model is constructed at each observational site and each grid point in order to assimilate available scientific knowledge about precipitation into the statistical model. The model is applied to two data sets on extreme precipitation, one uncorrected data set and one data set that is corrected for phase and wind. The observations are assumed to follow the generalized extreme value distribution. At the latent level, we implement SPDE spatial models for both the location and scale parameters of the likelihood. An efficient MCMC sampler which exploits the model structure is constructed, which yields fast continuous spatial predictions for spatially varying model parameters and quantiles.

Keywords. Extreme precipitation, latent Gaussian models, SPDE spatial models, MCMC block sampling

Address for correspondence: O.P. Geirsson, Department of Mathematics, School of Engineering and Natural Sciences, University of Iceland, Dunhagi 5, 107 Reykjavík, Iceland. E-mail: olipalli@gmail.com

1 Introduction

Extreme rainfall in Reykjavík on the 16th of August 1991 resulted in overloaded local drainage systems causing severe damage to industrial buildings, houses and apartments. Even though extreme precipitation events are rare, understanding their frequency and intensity is important for public safety and various types of long term agricultural, industrial and urban planning.

Meteorological models, of various complexities, have been proposed to model precipitation. For mountainous regions, such as much of Iceland, meteorological models that take into account the effects of topography as well as atmospheric information on orographic precipitation are needed. To that extent, Smith and Barstad [2004] proposed a quasi-analytic linear orographic method to simulate precipitation over a spatial domain which Crochet et al. [2007] adopted to precipitation in Iceland. The resulting model simulates daily precipitation on a 1 km by 1 km regular grid across Iceland over the years 1958-2002.

Although meteorological models are well suited to give insight into the spatial mean behavior of precipitation, they tend to deviate significantly from observations when predicting extreme precipitation. The first goal of this paper is to propose a statistical method that takes into account observations, spatial variation of the extremes and outputs from regional meteorological models to increase predictive power of extreme events. To that extent, a covariate based on the meteorological model of Crochet et al. [2007] is constructed in order to assimilate much of the scientific knowledge about precipitation into a statistical model.

In recent years, various statistical models have been proposed for modeling extreme precipitation, where models based on the generalized extreme value distribution are standard in the literature, see for example Sang and Gelfand [2009]. Furthermore, the Bayesian approach is well suited to quantify uncertainty of the underlying physical processes as further argued in [Tebaldi and Sansó, 2009]. In particular, the spatial variation can be modeled through the likelihood parameters as presented in e.g. [Davison et al., 2012, Hrafnkelsson et al., 2012]. Alternative modeling approaches have been explored, such as the peek over threshold methods with Generalized Pareto Distribution [Cooley et al., 2007].

Latent Gaussian models (LGMs) [Rue and Held, 2005], which form a flexible and practical subclass of Bayesian hierarchical models, play a dominant role in the vast literature on spatial statistics [Delfiner et al., 2009, Diggle et al., 1998, Guttorp and Gneiting, 2006]. In the LGM framework, Gaussian fields (GFs) appear at the latent level of the hierarchical model. However, posterior inference for GFs becomes increasingly computationally demanding as data sets get larger. Gaussian fields can be approximated by Gaussian Markov random fields (GMRFs), which increases the speed of computation significantly [Rue, 2001]. Although GMRFs are computationally efficient they become difficult to parameterize in a spatial setting. In recent work by Lindgren et al. [2011], numerical approximations to GFs with Matérn covariance structure are presented. The method constructs an approximation of a stochastic partial differential equation (SPDE)

on a triangulated mesh. The approximate solutions can then be used to construct a GMRF representation of the desired GF on the mesh. This allows for continuous spatial predictions by choosing appropriate basis functions for the approximation.

The second goal of the paper is to present a computationally efficient Bayesian hierarchical spatial model in order to obtain distributional properties of extreme precipitation on a high resolution grid. To that extent, an LGM is presented where the observations are assumed to follow the generalized extreme value distribution. The spatial variations are modeled through the location and scale parameters of the likelihood with SPDE spatial models. An MCMC split sampler [Geirsson et al., 2014] is applied to the model structure, yielding an efficient inference scheme. Furthermore, the model structure can also be used to make spatial predictions for the model parameters and quantiles of extreme precipitation on the high resolution grid.

The paper is organized as follows. The data and the meteorological model are presented in Section 2, where a method to construct covariates from outputs of the meteorological model is also outlined. Detailed description of the model structure, the inference method and spatial prediction is given in Section 3. Posterior results are presented and discussed in Section 4. The paper concludes with a discussion in Section 5.

2 The data

2.1 Observations

The observed data on precipitation were provided by the Icelandic Meteorological Office (IMO). Two types of data sets were explored in this paper, referred to as uncorrected and corrected data sets. The uncorrected data set contains raw observations of 24 hour annual maximum precipitation from 86 sites in Iceland over the years 1958 to 2006, as seen in [Crochet et al., 2007]. The latter data set is based on observations that have been corrected according to the dynamic correction model proposed by [F rland and Institutt, 1996]. The correction accounts for trace; wetting and evaporation losses; and for the catch deficiencies due to aerodynamic effects. Crochet et al. [2007] made corrections for 40 observational sites across Iceland over the years 1958 to 2006 on a daily basis, according to the phase (liquid, mixed, solid) of hydrometeors and the average daily wind speed and temperature recorded at each site. The observational sites for both data sets can be seen on Figure 1, where the uncorrected and corrected data sets are on the left and right panels, respectively.

Time series from four observational sites, Reykjav k,    ey, Akureyri and Kv sker, are shown on Figure 2 for the uncorrected and corrected data sets. Reykjav k and Akureyri were chosen because they are the most populated areas in Iceland.    ey was chosen to demonstrate the difference between the uncorrected and corrected data sets. Finally, Kv sker was chosen as it has the highest observed precipitation, and is by nature difficult to infer as it exhibits higher extremal behaviour than its neighbours.

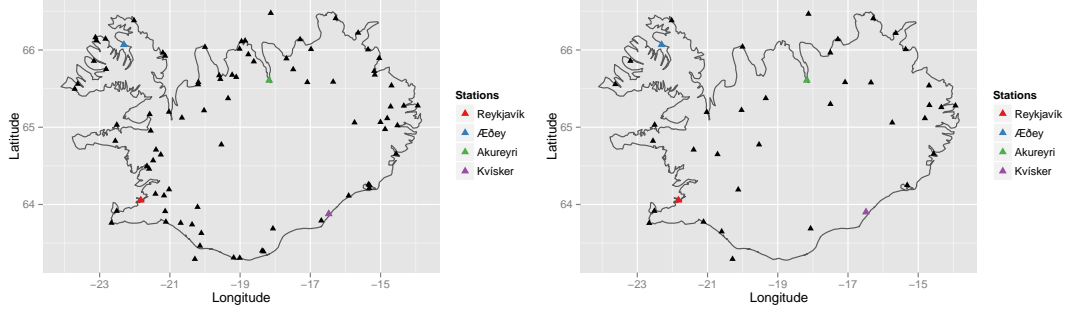


Figure 1: Observational sites. Left and right panel show the locations in the uncorrected and corrected data sets respectively

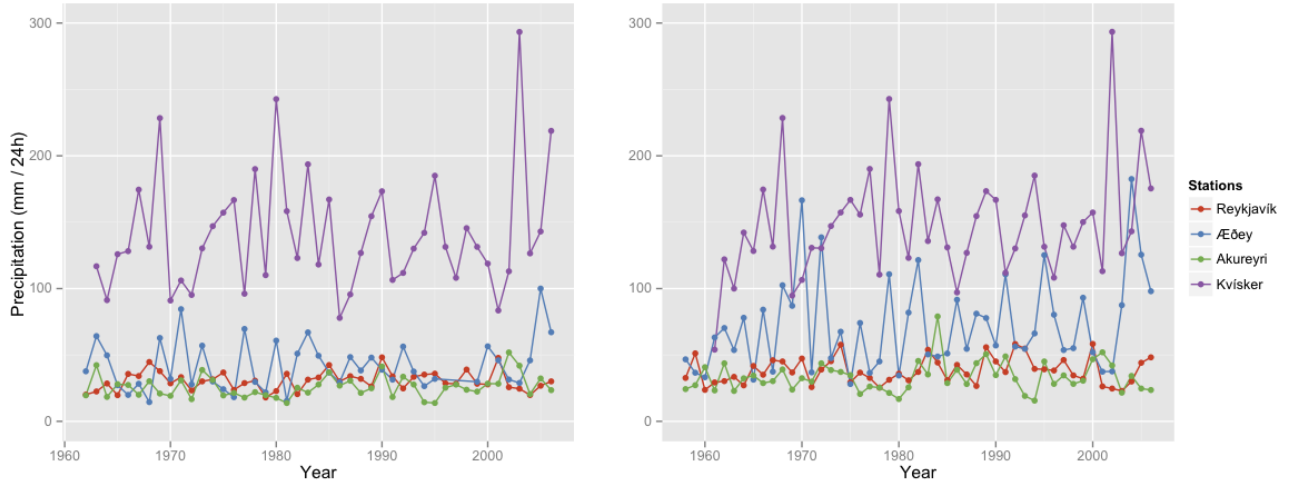


Figure 2: Times series. Left and right panel show the time series in the uncorrected and corrected data sets respectively

2.2 The meteorological model

The meteorological model that is used in this paper to improve spatial predictive power, is based on a linear orographic precipitation model proposed by Smith and Barstad [2004]. Crochet et al. [2007] have adapted the method to precipitation in Iceland. The model is driven by coarse resolution precipitation, wind and temperature data obtained from re-analyses (1958-2001) (ERA-40) [Uppala et al., 2005] and analyses (2002-2004) made by the European Center for Medium range Weather Forecast. The model takes into account the topography of the spatial domain, airflow dynamics, condensed water advection and downslope evaporation. This means that outputs from the model contains information about the underlying physical processes of precipitation. Note that, three free parameters of the meteorological models were calibrated using five years worth of daily precipitation data at forty observational sites across Iceland and precipitation information from three ice caps.

The resulting model simulates daily precipitation on a 1 km by 1 km regular grid across Iceland for the years 1958-2002. The study domain is on a $521 \text{ km} \times 361 \text{ km}$ regular grid. Thus, the simulated data set is (roughly) of the magnitude $3 \cdot 10^9$ data points, as the spatial dimension is 521×361 and the temporal dimension is 365×44 .

2.3 Covariates

Reasonable covariates for extreme precipitation should assimilate available spatial information such as the topography of the domain and the underlying physical processes of precipitation in order to increase the predictive power of the spatial predictions. By constructing covariates based on the meteorological model, the information about the above factors can be assimilated. Furthermore, Benestad et al. [2012] suggested that observed mean values of precipitation have high predictive power for extreme precipitation. Extending that argument, we propose that the calculated sample mean values from the meteorological model will have similar predictive power to the observed means. Moreover, the simulated data contain knowledge of the underlying physical processes. This leads to a high-quality, data consistent predicted mean over the entire spatial domain. This is a novel extension of the concept of Benestad et al. [2012] to spatial prediction of extremes. The means based on the meteorological model are referred to hereafter as simulated means.

The simulated means can be calculated at every grid point, which yields a $521 \text{ km} \times 361 \text{ km}$ regular grid of covariates. However, observational sites are not necessarily at the regular grid points. In order to construct the covariates at the observational sites we implemented the following spatial smoother.

Let \mathcal{G} denote the set of every regular grid point and let \mathcal{S} denote the set of the observational sites. Furthermore, let \bar{x}_j^* denote the simulated means at every grid point $j \in \mathcal{G}$ and \bar{x}_i denote observed means at each observational site $i \in \mathcal{S}$. For a given tuning distance $r > 0$, find every grid point $j \in \mathcal{G}$ that is within distance r of point $i \in \mathcal{S}$ and denote that index set with $D(r, i)$. Illustrative figure can be seen in Figure 3.

The goal is to construct a covariate at site i that uses information from the grid

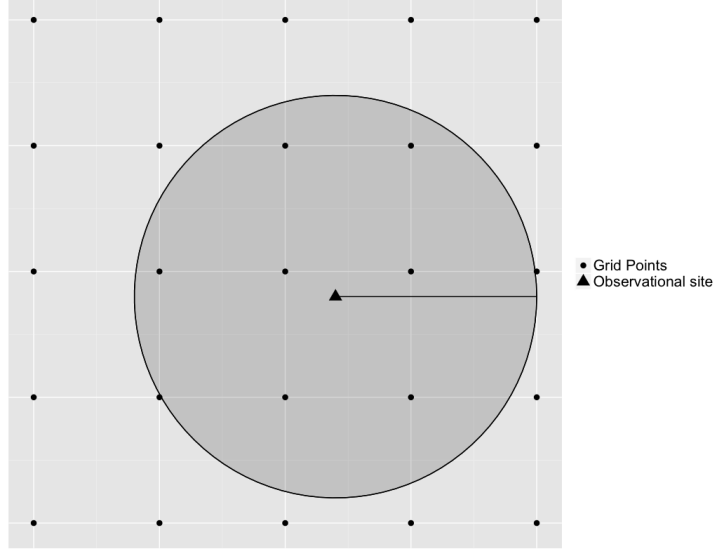


Figure 3: Illustrative figure, how the index set $D(r, i)$ is created for some radius r and observational site i . All the grid points within the circle then belong to $D(r, i)$

points in $D(r, i)$. Assign the following decay function

$$w_{ij} = 1 - F_{\alpha} \left(\frac{d_{ij}}{r} \right) \quad \text{for all } j \in D(r, i)$$

where F_{α} denotes the cumulative density function of the beta distribution with parameters α and α , and d_{ij} denotes the Euclidean distance from point i to point j . The two parameters of the beta distribution are chosen to be the same in order to avoid interactions between the tuning distance r and the shape of the decay function. The w_{ij} s describe the relative weights of the neighboring grid points in $D(r, i)$. The covariate at observation point i can then be constructed as the weighted mean of the surrounding simulated means, that is

$$\bar{x}_i^*(r, \alpha) = \frac{\sum_{j \in D(r, i)} w_{ij} \bar{x}_j^*}{\sum_{j \in D(r, i)} w_{ij}} \quad \text{for every } i \text{ in } \mathcal{S}. \quad (2.1)$$

Following the [Benestad et al., 2012] argument, the parameters in the spatial smoother are tuned such that the simulated means at the observational sites are close to the observed means. We chose to measure the distance between the simulated means, which are calculated with equation (2.1), and the observed means by using mean square distance. Then the optimal r and α can be found in the mean square sense by calculating the following

$$(r_0, \alpha_0) = \arg \min_{r, \alpha} \sum_{i \in \mathcal{S}} \|\bar{x}_i^*(r, \alpha) - \bar{x}_i\|^2.$$

The covariates that will be used in this paper are then $\bar{x}_i^* = \bar{x}_i^*(r_0, \alpha_0)$ for all observational sites $i \in \mathcal{S}$.

3 The model and inference

Guttorp and Gneiting [2006] proposed that using Gaussian fields with a Matérn covariance structure, called Matérn fields, is a flexible and interpretable way of modeling underlying physical processes of natural phenomena in a spatial setting. However, the resulting covariance matrices are dense and thus become computationally demanding in posterior inference and spatial predictions as data sets get larger.

Approximating Gaussian fields as Gaussian Markov random fields (GMRF) [Knorr-Held and Rue, 2002] provides an efficient framework for computationally efficient Gaussian models. GMRFs are parameterized with precision matrices, which are defined as inverses of covariance matrices. Although GMRFs have very good computational properties, there was no standard way to parametrize the precision matrix of a GMRF to achieve a predefined spatial covariance structure between two sites until Lindgren et al. [2011] addressed the issue. They proposed a method which computes a numerical approximation to the Matérn field on a triangulated mesh over the spatial domain, based a stochastic partial differential equation (SPDE) representation. The approximate solution $x(\mathbf{s})$, for each point \mathbf{s} in the spatial domain, is on the form

$$x(\mathbf{s}) = \sum_{k=1}^n \psi_k(\mathbf{s}_k) w_k \quad (3.1)$$

where n is the number of vertices in the mesh, ψ_k are piecewise linear basis functions and w_k are Gaussian weights. The approximate solutions can be used to construct a GMRF representation of the Matérn field at every point within the triangulated mesh.

In order to model the behavior of the underlying physical processes of extreme precipitation, the location and scale parameters in the likelihood are allowed to vary in space. To that extent, SPDE spatial models are used at the latent level of the proposed hierarchical model to describe continuously the spatial variation of these two latent parameters. Posterior inference and spatial predictions based on this approach are presented.

3.1 Model structure

The data are modeled with an LGM assuming the generalized extreme value distribution for the observations. That is, let y_{it} denote the annual maximum 24 hour precipitation at station i at year t , with a cumulative density function of the form

$$F(y_{it}) = \exp \left\{ - \left(1 + \xi \left(\frac{y_{it} - \mu_i}{\sigma_i} \right) \right)^{-1/\xi} \right\}, \quad i = 1, \dots, J, \quad t = 1, \dots, T$$

if $1 + \xi(x - \mu)/\sigma > 0$, $F(y_{it}) = 0$ otherwise. The parameters μ_i , σ_i and ξ are location, scale and shape parameters; J is the number of observational sites and T is the number of

years. These distributional assumptions are reasonable as the generalized extreme value distribution belongs to the family of extremal distributions and have desired asymptotic properties [Coles et al., 2001]. The observations are assumed to be independent conditioned on latent parameters. The conditional independence assumption implies that realizations of the spatial maxima will be everywhere discontinuous. This is somewhat unrealistic and may affect the width of the posterior intervals. On the other hand, most sites are far enough apart that they can be considered conditionally independent. Furthermore, the main purpose is to give spatial predictions of marginal quantiles using the SPDE approach, but not to simulate spatial realizations of the extreme precipitation for the next year or unobserved sites for an observed year. Hence, models with data-level dependence are not considered in this paper.

At the latent level of the model, we implement the following spatial model structure for the location parameter $\boldsymbol{\mu} = (\mu_1, \dots, \mu_J)$,

$$\boldsymbol{\mu} = \mathbf{X}_\mu \boldsymbol{\beta}_\mu + \mathbf{A}_\mathcal{S} \mathbf{u}_\mu + \mathbf{v}_\mu, \quad (3.2)$$

where \mathbf{X}_μ is a design matrix consisting of a vector of ones and the covariates \bar{x}_i^* that are based on the meteorological model; $\boldsymbol{\beta}_\mu$ are the corresponding weights; \mathbf{u}_μ is a spatial effect on a triangulated mesh; $\mathbf{A}_\mathcal{S}$ is a projection matrix; the matrix product $\mathbf{A}_\mathcal{S} \mathbf{u}_\mu$ then denotes the spatial effect at the observational sites, which captures the spatial variation in the data that is unexplained by the covariate and \mathbf{v}_μ is an unstructured random effect.

The spatial effect \mathbf{u}_μ is constructed using the SPDE approach. In order to obtain the SPDE structure for \mathbf{u}_μ , the domain of interest is subdivided into a mesh of non-intersecting triangles, see Figure 4. The GMRF representation is then constructed for the weights of the basis functions in (3.1) with a precision matrix $\mathbf{Q}_{u\mu}(\boldsymbol{\psi}_{u\mu})$ constructed with the SPDE approach of Lindgren et al. [2011]. The precision matrix is based on the geometry of the mesh and is, by construction, a sparse matrix. In this paper, we chose the smoothness parameter to be $\nu = 1$, which corresponds to an almost once differentiable Matérn field. The precision matrix $\mathbf{Q}_{u\mu}(\boldsymbol{\psi}_{u\mu})$ has two parameters, $\boldsymbol{\psi}_{u\mu} = (\kappa_{u\mu}, \omega_{u\mu})$, which enter the model as hyperparameters. The hyperparameter $\kappa_{u\mu}$ is inversely proportional to the range of the approximate Matérn field and the hyperparameter $\omega_{u\mu}$ is related to the marginal variance.

The locations of \mathbf{u}_μ in space are on the vertices of the mesh, which are not necessarily at the observational sites. However, since the approximate representation (3.1) is assumed to have piecewise linear basis functions, an approximate solution can also be found within each triangle as a convex linear combination of the approximate solutions at the surrounding vertices. The spatial effect \mathbf{u}_μ can be projected linearly onto every point within every triangle of the mesh. The matrix $\mathbf{A}_\mathcal{S}$ denotes the linear projection from the mesh vertices onto the observational sites. Note that the number of non-zero entries in each line of $\mathbf{A}_\mathcal{S}$ is three and the sum of every line is one.

Analogous spatial structure is also implemented for the scale parameter σ_i on a logarithmic scale. That is, let $\tau_i = \log \sigma_i$ and then model $\boldsymbol{\tau} = (\tau_1, \dots, \tau_J)$ as

$$\boldsymbol{\tau} = \mathbf{X}_\tau \boldsymbol{\beta}_\tau + \mathbf{A}_\mathcal{S} \mathbf{u}_\tau + \mathbf{v}_\tau.$$

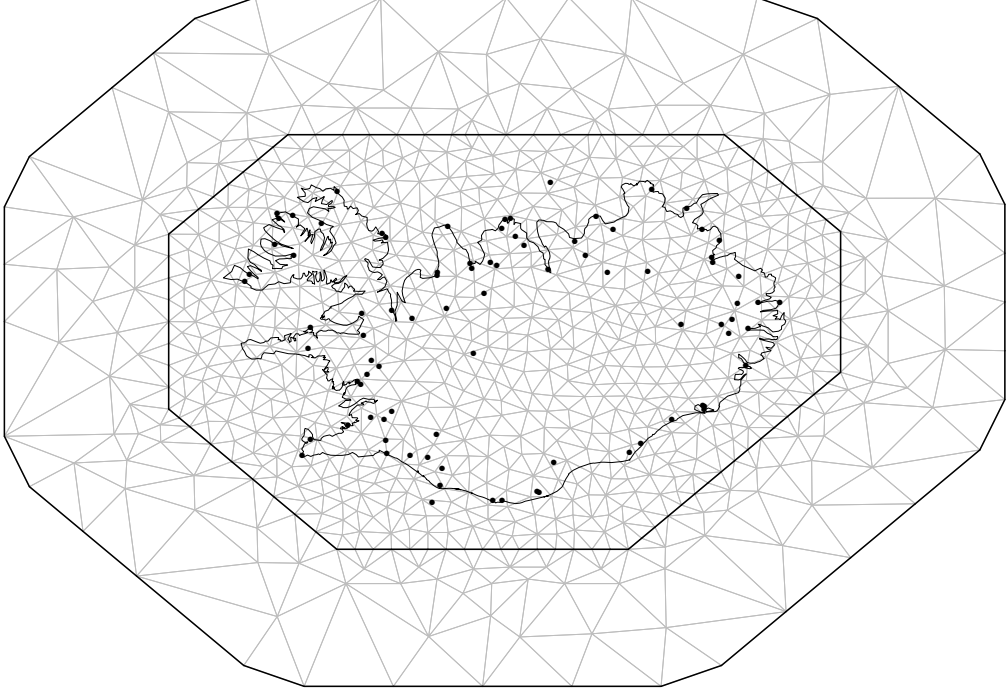


Figure 4: Triangulated mesh over Iceland

However, the design matrix \mathbf{X}_τ consists of a vector of ones and a covariate based on the meteorological model on a logarithmic scale, that is $\log \bar{x}_i^*$.

3.2 Prior selection

The following prior distributions were assigned for the latent parameters,

$$\begin{aligned} \beta_\mu &\sim \mathcal{N}(\mathbf{0}, \kappa_{\beta\mu}^{-1} \mathbf{I}), & \beta_\tau &\sim \mathcal{N}(\mathbf{0}, \kappa_{\beta\tau}^{-1} \mathbf{I}), \\ \mathbf{u}_\mu &\sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{u\mu}^{-1}(\psi_{u\mu})), & \mathbf{u}_\tau &\sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{u\tau}^{-1}(\psi_{u\tau})), \\ \mathbf{v}_\mu &\sim \mathcal{N}(\mathbf{0}, \kappa_{v\mu}^{-1} \mathbf{I}), & \mathbf{v}_\tau &\sim \mathcal{N}(\mathbf{0}, \kappa_{v\tau}^{-1} \mathbf{I}). \end{aligned}$$

The parameters β_μ and β_τ are assumed *a priori* to have a low precision on their native scales in order to let the data play the dominate role in their inference. Thus, the parameter values $\kappa_{\beta\mu} = 0.0025$ and $\kappa_{\beta\tau} = 0.25$ were chosen for the prior distributions.

Lognormal prior distributions with fixed parameters were assigned to the hyperparameters of the spatial fields \mathbf{u}_μ and \mathbf{u}_τ , that is

$$\begin{aligned} \kappa_{u\mu} &\sim \mathcal{LN}(\mu_{\kappa u\mu}, \sigma_{\mu\kappa}^2), & \kappa_{u\tau} &\sim \mathcal{LN}(\mu_{\tau\kappa}, \sigma_{\tau\kappa}^2) \\ \omega_{u\mu} &\sim \mathcal{LN}(\mu_{\mu\omega}, \sigma_{\mu\omega}^2), & \omega_{u\tau} &\sim \mathcal{LN}(\mu_{\tau\omega}, \sigma_{\tau\omega}^2) \end{aligned}$$

where \mathcal{LN} denotes the lognormal distribution. The fixed parameters in the priors for $\kappa_{u\mu}$ and $\kappa_{u\tau}$ were chosen in a weakly informative manner to ensure that the range of

the spatial effects is sensible relative to the size of domain of interest. The relationship between the parameters $\kappa_{u\mu}$ and $\kappa_{u\tau}$ and the spatial range of the approximated Matérn field used in this paper is $\rho = \sqrt{8\nu}/\kappa$, which corresponds to correlation near 0.1 at a distance ρ . Consequently, the prior distributions for $\kappa_{u\mu}$ and $\kappa_{u\tau}$ were chosen such that the corresponding ρ 's range up to half of the length of the domain, which in this paper is approximately 150 km. To that extent, the values $\mu_{\mu\kappa} = -2.7$, $\sigma_{\mu\kappa} = 0.45$, $\mu_{\tau\kappa} = -2.5$, $\sigma_{\tau\kappa} = 0.45$ were chosen.

Site specific maximum likelihood estimates for $\boldsymbol{\mu}$ and $\boldsymbol{\tau}$ in an exploratory data analysis and the relation to their corresponding covariates \mathbf{X}_μ and \mathbf{X}_τ suggest that the marginal standard deviations for \mathbf{u}_μ and \mathbf{u}_τ should exceed 50 and 5, respectively, with low probability on their native scales. The relationship between the parameters $\omega_{u\mu}$ and $\omega_{u\tau}$ and the marginal standard deviation of the Matérn field is approximately $1/(\sqrt{4\pi\kappa\omega})$. Which led to the parameter values $\mu_{\mu\omega} = -1.1$, $\sigma_{\mu\omega} = 0.1$, $\mu_{\tau\omega} = 1.8$, $\sigma_{\tau\omega} = 0.1$.

The unstructured random effects \mathbf{v}_μ and \mathbf{v}_τ capture the variation in the data which is unexplained by the covariates and the spatial models. The following prior distributions for the precision parameters of \mathbf{v}_μ and \mathbf{v}_τ were assigned

$$\kappa_{v\mu} \sim \mathcal{LN}(\mu_{\mu v}, \sigma_{\mu v}^2), \quad \kappa_{v\tau} \sim \mathcal{LN}(\mu_{\tau v}, \sigma_{\tau v}^2).$$

The standard deviations of \mathbf{v}_μ and \mathbf{v}_τ are *a priori* believed to be mainly between 2 and 20, and 0.1 and 1.5, respectively. Hence, the values $\mu_{\mu v} = -4.5$, $\sigma_{\mu v} = 0.45$, $\mu_{\tau v} = 1.5$, and $\sigma_{\tau v} = 0.1$ were chosen.

Plots of the prior distributions for all the hyperparameters $\kappa_{u\mu}$, $\kappa_{u\tau}$, $\omega_{u\mu}$, $\omega_{u\tau}$, $\kappa_{v\mu}$ and $\kappa_{v\tau}$ are shown on transformed scales for interpretability. That is, the range of the two spatial fields are shown in Figures 6a and 6b, which are functions of $\kappa_{u\mu}$ and $\kappa_{u\tau}$; the marginal standard deviation of the spatial fields are shown in Figures 6c and 6d, which are functions of $\omega_{u\mu}$ and $\omega_{u\tau}$; and standard deviation of the unstructured random effects are shown in Figures 6e and 6f, which are functions of $\kappa_{v\mu}$ and $\kappa_{v\tau}$.

Due to lack of data and for illustration purposes, the shape parameter ξ was assumed constant as opposed to having it varying over the spatial domain. Thus, the shape parameters was assigned the prior distribution

$$\xi \sim \mathcal{N}(0, \kappa_\xi^{-1})$$

where $\kappa_\xi = 2$. Note that ξ is inferred in the model as a latent parameter as seen in Hrafnkelsson et al. [2012].

3.3 Posterior inference

Due to the proposed model structure, in particular the spatial model on the scale parameters, MCMC methods were necessary to make posterior inference as opposed to approximation methods such as INLA [Rue et al., 2009]. However, standard MCMC methods like single site updating converged slowly and mixed poorly as many model parameters were heavily correlated in the posterior. To address this issue, posterior inference was done by using the MCMC split sampler Geirsson et al. [2014].

In order to implement the MCMC split sampler, the model parameters are set up as follows. First, define the data-rich part $\boldsymbol{\eta} = (\boldsymbol{\mu}, \boldsymbol{\tau}, \xi)$ and the data-poor part $\boldsymbol{\nu} = (\boldsymbol{\beta}_\mu, \mathbf{u}_\mu, \boldsymbol{\beta}_\tau, \mathbf{u}_\tau)$. The latent field of the proposed LGM can then be written as

$$\mathbf{x} = (\boldsymbol{\eta}, \boldsymbol{\nu}) = (\boldsymbol{\mu}, \boldsymbol{\tau}, \xi, \boldsymbol{\beta}_\mu, \mathbf{u}_\mu, \boldsymbol{\beta}_\tau, \mathbf{u}_\tau)$$

with hyperparameters, $\boldsymbol{\theta} = (\kappa_{u\mu}, \omega_{u\mu}, \kappa_{u\tau}, \omega_{u\tau}, \kappa_{v\mu}, \kappa_{v\tau})$. Then define

$$\mathbf{Z} = \begin{pmatrix} \mathbf{X}_\mu & \mathbf{A}_S & \cdot & \cdot \\ \cdot & \cdot & \mathbf{X}_\tau & \mathbf{A}_S \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \quad \mathbf{Q}_\epsilon = \begin{pmatrix} \kappa_{v\mu} \mathbf{I} & \cdot & \cdot \\ \cdot & \kappa_{v\tau} \mathbf{I} & \cdot \\ \cdot & \cdot & \kappa_\xi \end{pmatrix}, \quad \mathbf{Q}_\nu = \begin{pmatrix} \kappa_{\beta\mu} \mathbf{I} & \cdot & \cdot & \cdot \\ \cdot & \mathbf{Q}_{u\mu} & \cdot & \cdot \\ \cdot & \cdot & \kappa_{\beta\tau} \mathbf{I} & \cdot \\ \cdot & \cdot & \cdot & \mathbf{Q}_{u\tau} \end{pmatrix}.$$

where we use $\mathbf{Q}_{u\mu}$ instead of $\mathbf{Q}_{u\mu}(\psi_{u\mu})$ for notational convenience and the dotted entries denote zero elements. The joint prior distribution of the latent field conditioned on the hyperparameters, i.e. $\mathbf{x}|\boldsymbol{\theta}$, is then a mean zero Gaussian with the following precision matrix

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_\epsilon & -\mathbf{Q}_\epsilon \mathbf{Z} \\ -\mathbf{Z}^\top \mathbf{Q}_\epsilon & \mathbf{Q}_\nu + \mathbf{Z}^\top \mathbf{Q}_\epsilon \mathbf{Z} \end{pmatrix}.$$

The corresponding posterior density is,

$$\pi(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y}) = \pi(\boldsymbol{\eta}, \boldsymbol{\nu}, \boldsymbol{\theta}|\mathbf{y}) \propto \pi(\mathbf{y}|\boldsymbol{\eta})\pi(\boldsymbol{\eta}, \boldsymbol{\nu}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$$

A sample from the posterior density $\pi(\boldsymbol{\eta}, \boldsymbol{\nu}, \boldsymbol{\theta}|\mathbf{y})$ is then obtained in with a two block Gibbs sampler type, and is briefly outlined below. The details can be seen in Geirsson et al. [2014].

3.3.1 Data-rich part

The structure of the data-rich part is as follows. Let $\boldsymbol{\eta}_A = (\boldsymbol{\mu}, \boldsymbol{\tau})$. In order to exploit the conditionally independence structure of the model, we sample from $\pi(\boldsymbol{\eta}_A|\mathbf{y}, \xi, \boldsymbol{\nu}, \boldsymbol{\theta})$ and $\pi(\xi|\mathbf{y}, \boldsymbol{\eta}_A, \boldsymbol{\nu}, \boldsymbol{\theta})$ in two blocks. To sample from $\pi(\boldsymbol{\eta}_A|\mathbf{y}, \xi, \boldsymbol{\nu}, \boldsymbol{\theta})$, start by defining

$$\mathbf{Z}_A = \begin{pmatrix} \mathbf{X}_\mu & \mathbf{A}_S & \cdot & \cdot \\ \cdot & \cdot & \mathbf{X}_\tau & \mathbf{A}_S \end{pmatrix}, \quad \mathbf{Q}_A = \begin{pmatrix} \kappa_{v\mu} \mathbf{I} & \cdot \\ \cdot & \kappa_{v\tau} \mathbf{I} \end{pmatrix}$$

and

$$\boldsymbol{\eta}_{A(i)} = (\mu_i, \tau_i) \quad \text{for } i = 1, \dots, J.$$

The logarithm of the conditional posterior becomes

$$\log \pi(\boldsymbol{\eta}_A|\mathbf{y}, \xi, \boldsymbol{\nu}, \boldsymbol{\theta}) = -\frac{1}{2} \boldsymbol{\eta}_A^\top \mathbf{Q}_A \boldsymbol{\eta}_A + (\mathbf{Q}_A \mathbf{Z} \boldsymbol{\nu})^\top \boldsymbol{\eta}_A + f_A(\boldsymbol{\eta}_A) + \text{const}$$

where

$$f_A(\boldsymbol{\eta}_A) = \sum_{i=1}^J f_i(\boldsymbol{\eta}_{A(i)}) = \sum_{i=1}^J \sum_{t \in \mathcal{A}_i} \log f_{\text{gev}}(y_{it} | \mu_i, \exp \tau_i, \xi),$$

where f_{gev} denotes the density of the generalized extreme value distribution and the set \mathcal{A}_i contains the indices of the years t observed at site i . The steps of the sampler are outlined in Algorithm 1.

Algorithm 1 Obtain the $(k+1)$ -th sample from $\pi(\boldsymbol{\eta}_A | \mathbf{y}, \xi, \boldsymbol{\nu}, \boldsymbol{\theta})$

Input: $\boldsymbol{\eta}_A^k$

- 1: Find the mode $\boldsymbol{\eta}_A^0 = \arg \max_{\boldsymbol{\eta}_A} \log \pi(\boldsymbol{\eta}_A | \mathbf{y}, \xi, \boldsymbol{\nu}, \boldsymbol{\theta})$
 - 2: Calculate $\mathbf{H} = \nabla^2 f_A(\boldsymbol{\eta}_A^0)$ and $\mathbf{b} = \nabla f_A(\boldsymbol{\eta}_A^0) - \mathbf{H} \boldsymbol{\eta}_A^0$
 - 3: Sample $\boldsymbol{\eta}_A^* \sim \mathcal{N}(\boldsymbol{\eta}_A^0, (\mathbf{Q}_A - \mathbf{H})^{-1})$
 - 4: Calculate $\boldsymbol{\rho}(\boldsymbol{\eta}^k)$ and $\boldsymbol{\rho}(\boldsymbol{\eta}^*)$, where $\boldsymbol{\rho}(\boldsymbol{\eta}) = (-\frac{1}{2} \boldsymbol{\eta}^\top \mathbf{H} - \mathbf{b}^\top) \circ \boldsymbol{\eta}$ and \circ denotes an entrywise product
 - 5: **for** $i = 1, \dots, J$
 - 6: Calculate $r_i = f_i(\boldsymbol{\eta}_{A(i)}^*) + \boldsymbol{\rho}(\boldsymbol{\eta}^*)_{(i)}^\top \mathbf{1} - (f_i(\boldsymbol{\eta}_{A(i)}^k) + \boldsymbol{\rho}(\boldsymbol{\eta}^k)_{(i)}^\top \mathbf{1})$
 - 7: Calculate $\alpha_i = \min\{1, \exp r_i\}$
 - 8: Sample $u_i \sim \mathcal{U}(0, 1)$
 - 9: **if** $\alpha_i > u_i$
 - 10: $\boldsymbol{\eta}_{A(i)}^{k+1} = \boldsymbol{\eta}_{A(i)}^*$
 - 11: **else if** $\alpha_i < u_i$
 - 12: $\boldsymbol{\eta}_{A(i)}^{k+1} = \boldsymbol{\eta}_{A(i)}^k$
 - 13: **end if**
 - 14: **end for**
- Output:** $\boldsymbol{\eta}^{k+1}$
-

In order to sample from $\pi(\xi | \mathbf{y}, \boldsymbol{\eta}_A, \boldsymbol{\nu}, \boldsymbol{\theta})$, note that the logarithm of the corresponding conditional posterior is

$$\log \pi(\xi | \mathbf{y}, \boldsymbol{\eta}_A, \boldsymbol{\nu}, \boldsymbol{\theta}) = \log \pi(\xi) + f_B(\xi)$$

where

$$f_B(\xi) = \sum_{i=1}^J \sum_{t \in \mathcal{A}_i} \log f_{\text{gev}}(y_{it} | \mu_i, \exp \tau_i, \xi),$$

The steps of the sampler are outlined in Algorithm 2.

3.3.2 Data-poor part

The proposal strategy suggested in [Knorr-Held and Rue, 2002] is used for each element of $\boldsymbol{\theta}$. That is, let $\theta_i^* = f \theta_i^k$ where the scaling factor f has the density

$$\pi(f) \propto 1 + 1/f \quad \text{for } f \in [1/F, F] \quad (3.3)$$

Algorithm 2 Obtain the $(k + 1)$ -th sample from $\pi(\xi|\mathbf{y}, \boldsymbol{\eta}_A, \boldsymbol{\nu}, \boldsymbol{\theta})$

Input: ξ^k

- 1: Find the mode $\xi^0 = \arg \max_{\xi} \log \pi(\xi|\mathbf{y}, \boldsymbol{\eta}_A, \boldsymbol{\nu}, \boldsymbol{\theta})$
- 2: Sample $\xi^* \sim \mathcal{N}(\xi^0, (\kappa_{\xi} - f''(\xi^0))^{-1})$
- 3: Calculate $\rho(\xi^k)$ and $\rho(\xi^*)$, where $\rho(\xi) = \left(-\frac{f''(\xi^0)}{2}\xi - \left(f'(\xi^0) - f''(\xi^0)\xi^0\right)\right)\xi$
- 4: Calculate $r = f_B(\xi^*) + \rho(\xi^*) - (f_B(\xi^k) + \rho(\xi^k))$
- 5: Calculate $\alpha = \min\{1, \exp r\}$
- 6: Sample $u \sim \mathcal{U}(0, 1)$
- 7: **if** $\alpha > u$
- 8: $\xi^{k+1} = \xi^*$
- 9: **else if** $\alpha < u$
- 10: $\xi^{k+1} = \xi^k$
- 11: **end if**

Output: ξ^{k+1}

where $F > 1$ is a tuning parameter. It can be shown that this a symmetric proposal density in the sense that

$$q(\theta_i^*|\theta_i^k) = q(\theta_i^k|\theta_i^*)$$

A joint proposal strategy for $(\boldsymbol{\nu}, \boldsymbol{\theta})$ is implemented as seen in Geirsson et al. [2014]. The steps of the sampler are outlined in Algorithm 3.

3.4 Spatial prediction

By using the SPDE spatial model structure, the posterior distribution of all the spatially varying model parameters can be obtained at every regular grid point in \mathcal{G} . The details are as follows.

Let $\mathbf{u}^{[k]}$ be the k -th posterior MCMC sample of the spatial effects, for either the location parameter or the scale parameter. The spatial effects $\mathbf{u}^{[k]}$ are located at the vertices of the triangles in the mesh, which do not necessarily coincide with the regular grid points in \mathcal{G} . However, since every regular grid point belongs to some triangle in the mesh, the k -th posterior MCMC sample for the spatial effects at the regular grid points can be obtained with a convex linear combination of $\mathbf{u}^{[k]}$, that is,

$$\mathbf{u}_{\mathcal{G}}^{[k]} = \mathbf{A}_{\mathcal{G}}\mathbf{u}^{[k]}. \quad (3.4)$$

where the matrix $\mathbf{A}_{\mathcal{G}}$ denotes the linear projection from the vertices of the mesh onto the regular grid points in \mathcal{G} . The term $\mathbf{u}_{\mathcal{G}}^{[k]}$ then serves as the k -th posterior MCMC sample for the spatial effects on the regular grid \mathcal{G} , which is calculated in post calculations after the MCMC run. Therefore, after calculating $\mathbf{u}_{\mathcal{G}}^{[k]}$ for every iteration k , posterior statistics for the spatial effects can be obtained for every point in \mathcal{G} , in particular posterior means and standard deviations.

Algorithm 3 Obtain the $(k + 1)$ -th sample from $\pi(\boldsymbol{\nu}, \boldsymbol{\theta} | \mathbf{y}, \boldsymbol{\eta})$

Input: $(\boldsymbol{\nu}^k, \boldsymbol{\theta}^k)$

- 1: Sample each element of $\boldsymbol{\theta}^*$ from the proposal density in (3.3)
- 2: Sample $\boldsymbol{\nu}^*$ from

$$\boldsymbol{\nu}^* | \boldsymbol{\eta}, \boldsymbol{\theta}^* \sim \mathcal{N} \left(\left(\mathbf{Q}_{\nu} + \mathbf{Z}^{\top} \mathbf{Q}_{\epsilon} \mathbf{Z} \right)^{-1} \mathbf{Z}^{\top} \mathbf{Q}_{\epsilon} \boldsymbol{\eta}, \left(\mathbf{Q}_{\nu} + \mathbf{Z}^{\top} \mathbf{Q}_{\epsilon} \mathbf{Z} \right)^{-1} \right)$$

- 3: Calculate

$$r = \frac{\pi(\boldsymbol{\theta}^*)}{\pi(\boldsymbol{\theta}^k)} \cdot \frac{\pi(\boldsymbol{\eta} | \boldsymbol{\nu}^*, \boldsymbol{\theta}^*) \pi(\boldsymbol{\nu}^* | \boldsymbol{\theta}^*)}{\pi(\boldsymbol{\nu}^* | \boldsymbol{\eta}, \boldsymbol{\theta}^*)} \cdot \frac{\pi(\boldsymbol{\nu}^k | \boldsymbol{\eta}, \boldsymbol{\theta}^k)}{\pi(\boldsymbol{\eta} | \boldsymbol{\nu}^k, \boldsymbol{\theta}^k) \pi(\boldsymbol{\nu}^k | \boldsymbol{\theta}^k)}$$

- 4: Calculate $\alpha = \min \{1, r\}$
 - 5: Sample $u \sim \mathcal{U}(0, 1)$
 - 6: **if** $\alpha > u$
 - 7: $(\boldsymbol{\nu}^{k+1}, \boldsymbol{\theta}^{k+1}) = (\boldsymbol{\nu}^*, \boldsymbol{\theta}^*)$
 - 8: **else if** $\alpha < u$
 - 9: $(\boldsymbol{\nu}^{k+1}, \boldsymbol{\theta}^{k+1}) = (\boldsymbol{\nu}^k, \boldsymbol{\theta}^k)$
 - 10: **end if**
- Output:** $(\boldsymbol{\nu}^{k+1}, \boldsymbol{\theta}^{k+1})$
-

The covariates, for both the location and scale parameter, are available by construction at every regular grid point. Thus, the k -th posterior MCMC sample for the location parameter can be calculated with

$$\boldsymbol{\mu}_{\mathcal{G}}^{[k]} = \mathbf{X}_{\mu} \boldsymbol{\beta}^{[k]} + \mathbf{A}_{\mathcal{G}} \mathbf{u}_{\mu}^{[k]}. \quad (3.5)$$

for every regular grid point. Analogous results hold for the scale parameter. Consequently, the posterior distribution of the p -th quantile of the generalized extreme value distribution can be calculated at every regular grid point. The p -th quantile function of the generalized extreme value distribution is

$$q_p(\mu, \tau, \xi) = \mu + \frac{\exp(\tau)}{\xi} \left(-\log(p)^{-\xi} - 1 \right). \quad (3.6)$$

The k -th posterior MCMC sample for the p -th quantile is then obtained by plugging in the k -th posterior MCMC samples for the location, scale and shape parameters. Thus, posterior samples and posterior statistics for the p -th quantile are obtained for every point in \mathcal{G} .

4 Results

The main objective of this analysis is to obtain spatial predictions for the location and the log-scale parameters and to evaluate the posterior distribution of the 0.95 quantile of maximum precipitation across Iceland on the regular grid \mathcal{G} . In Section 4.1, results from the MCMC convergence diagnostics are briefly summarized while in Section 4.2 tables and figures of posterior estimates of the model parameters are given and discussed. In Section 4.3 figures of spatial predictions are shown and discussed.

4.1 Convergence diagnostics

The following results are based on four MCMC chains from the sampler in Section 3.3. Each chain is calculated with 35000 iterations where 10000 iterations were burned in. Runtime, on a modern desktop (Ivy Bridge Intel Core i7-3770K, 16GB RAM and solid state hard drive), was approximately four hours for each data set. All the calculations were done using R.

Gelman–Rubin statistics [Brooks and Gelman, 1998] were calculated for all model parameters based on the four MCMC chains for both data sets. The Gelman–Rubin statistics for all the model parameters in both data sets, were evaluated as approximately 1, indicating that all the MCMC chains have converged in the mean. Figure 11 shows Gelman–Rubin plots for the location parameter $\boldsymbol{\mu}$ and log-scale parameter $\boldsymbol{\tau}$ in Reykjavík, Æðey, Akureyri and Kvísker; and for the coefficients of the covariates $\boldsymbol{\beta}_{\mu}$ and $\boldsymbol{\beta}_{\tau}$. Furthermore, Figure 12 shows Gelman–Rubin plots for all the hyperparameters. Both figures are based on the uncorrected data set. Similar results hold for the corrected data set. Moreover, Figures 11 and 12 indicate that the sampler converges after 2000 iterations after burn-in.

Others convergence diagnostics plots can be seen in the Appendix for the same set of parameters, where the four MCMC chains were evaluated. Figure 13 shows trace and density plots of the location and log-scale parameters for the same observational sites as above; autocorrelation and running mean plots can be seen in Figure 14. Figure 15 and Figure 16 show the same two convergence diagnostics plots for the coefficients of the covariates and all the hyperparameters. The MCMC chains for the latent parameters and the coefficients of the covariates exhibit neglectable autocorrelation after lag 10. Samples of the hyperparameters show some autocorrelation, but within an acceptable range as they are highly correlated in the posterior.

These convergence diagnostics results demonstrate the sampler’s efficiency, and strengthen the claim that the sampler converges quickly. We state, with fair amounts of confidence, that the sampler is highly efficient for this model.

4.2 Posterior estimates

The following posterior estimates are based on the four MCMC chains after burn-in. Figure 5 shows the posterior mean with 95% posterior interval for both the location and log-scale parameters for each observational site. The sites are placed on the x -axis as follows. The leftmost site on the x -axis is in Reykjavík. The rest of sites are places on the x -axis corresponding to a clockwise labeling across the sites shown in Figure 1. The results in Figure 5 suggest that both the average and standard deviation of extreme precipitation is the highest and the south-eastern part in Iceland, and the lowest in the northern part.

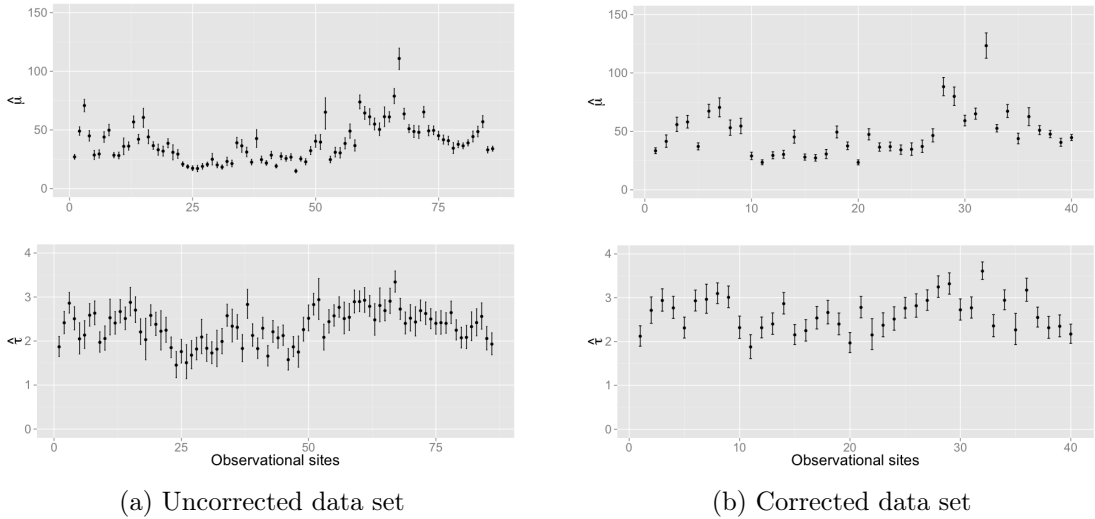


Figure 5: The first and second row show posterior estimates of μ and τ , respectively, for each observational site

Table 1 shows the posterior mean, posterior standard deviation and posterior 0.025 and 0.975 quantile estimates for the non-spatially varying model parameters. The pos-

terior results are similar for all the parameters between the two data sets, except for the covariate coefficients $\beta_{\mu 1}$ and $\beta_{\mu 2}$. That is to be expected, as every measurement in the corrected data set is greater or equal to the corresponding measurement in the uncorrected data set, by construction. The following results are discussed for the uncorrected data set. Analogous results and interpretation hold for the corrected data set.

Parameter	Uncorrected data				Corrected data			
	0.025	mean	0.975	sd	0.025	mean	0.975	sd
$\beta_{\mu 1}$	-4.65	2.41	8.96	3.46	2.53	11.76	20.82	4.66
$\beta_{\mu 2}$	9.14	10.97	12.87	0.95	9.04	11.59	14.24	1.33
$\beta_{\tau 1}$	1.24	1.52	1.79	0.14	1.68	2.10	2.52	0.21
$\beta_{\tau 2}$	0.49	0.71	0.94	0.12	0.13	0.50	0.87	0.19
ξ	0.06	0.09	0.12	0.02	0.08	0.12	0.15	0.02
$\omega_{u\mu}$	0.31	0.45	0.62	0.08	0.28	0.40	0.55	0.07
$\kappa_{u\mu}$	0.05	0.08	0.10	0.01	0.05	0.07	0.10	0.01
$\omega_{u\tau}$	5.42	5.97	6.57	0.29	5.37	5.93	6.53	0.29
$\kappa_{u\tau}$	0.11	0.13	0.17	0.02	0.09	0.12	0.16	0.02
$\kappa_{v\mu}$	0.01	0.02	0.04	0.01	0.01	0.02	0.03	0.01
$\kappa_{v\tau}$	7.85	12.61	19.05	2.87	4.04	7.10	11.61	1.94

Table 1: Posterior 0.025 quantile, mean, 0.975 quantile and standard deviation of the above model parameters

The posterior density for $\beta_{\mu 2}$ suggests that the effects of the constructed meteorological covariate is close to 11 and it is positive as its 95% posterior interval is above zero. Assuming the simulated means approximate the observed means well enough, these results indicate that the location parameter describing the extreme precipitation is roughly 11 times the corresponding annual mean in case of the uncorrected data.

The posterior density for $\beta_{\tau 1}$ and $\beta_{\tau 2}$ yield a point estimates close to 1.5 and 0.7, respectively, and it reveals that they are both positive as their 95% posterior intervals are above zero. That means that the relationship between the simulated means and the scale parameter of the generalized extreme value distribution can be roughly summarized as

$$\hat{\sigma}_i = e^{\hat{\beta}_1 \bar{x}_i^{*\hat{\beta}_2}} \approx 4.6 \cdot \bar{x}_i^{*0.7}$$

The estimate for the shape parameter 0.09 and its posterior interval is approximately (0.06, 0.12). That indicates that the proposed generalized extreme values distributions does not have an upper bound and its first and second moments are finite.

The posterior estimates of the parameters $\kappa_{u\mu}$ and $\kappa_{u\tau}$, suggest that the correlation between two points in space is near 0.1 at a roughly 40 km distance for the location parameter, and roughly 25 km distance for the scale parameter. The posterior estimates of the parameters $\omega_{u\mu}$ and $\omega_{u\tau}$, indicate that the marginal standard deviation of \mathbf{u}_μ is approximately 8 and 0.4 for \mathbf{u}_τ , which is the spatial effect for the scale parameter on a logarithmic scale.

The posterior estimates of $\kappa_{v\mu}$ indicate that the posterior standard deviation of the unstructured random effect \mathbf{v}_μ is approximately 7. Which in turn suggests that there is some variation left in the data that is unexplained by the covariates and the spatial model. That behavior might be due to the fact that some observational sites, like Kvísker, are located in areas with high topographical variation that yield a non-stationary behavior in the spatial field corresponding to the location parameter. Similar behavior is observed for the scale parameter, as the posterior estimates of $\kappa_{v\tau}$ indicate that the posterior standard deviation of the unstructured random effect \mathbf{v}_τ is approximately 0.28 on a logarithmic scale.

In Figure 6 prior distributions for all the hyperparameters are shown, along with corresponding posterior distributions based on the uncorrected data set.

4.3 Model evaluation

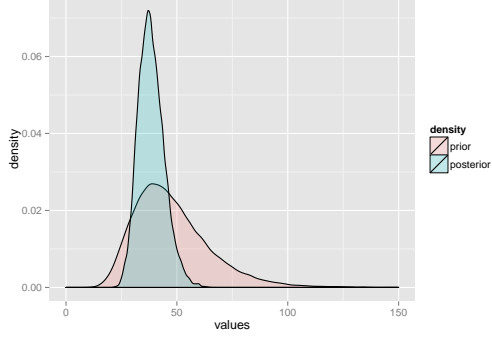
In Figure 7 empirical cumulative distributions for the observational sites Reykjavík, Aðey, Akureyri and Kvísker are compared to corresponding posterior cumulative distributions functions uncorrected data set. The results indicate that the model describes the data well. The observation at Kvísker deviate the most from the model in the uncorrected data set, while that behavior is not seen in the corrected data set. This might be due to that fact that there are observational sites close to Kvísker in the uncorrected data set that both show much lower extremal observations and are not included in the corrected data set.

An overall time effect for location parameter was evaluated based on the fitted values at each observational site. That is, let $m_{it} = y_{it} - E(y_{it})$, where $E(y_{it})$ is the expected value of the g.e.v. model with parameter values based on posterior means. A likelihood ratio test was applied to a model for the m 's with an overall time effect against a model for the m 's without a time effect. The results indicated no significant difference between the two models (significance level 0.05). Thus, there was no reason to add an overall time effect to the proposed model.

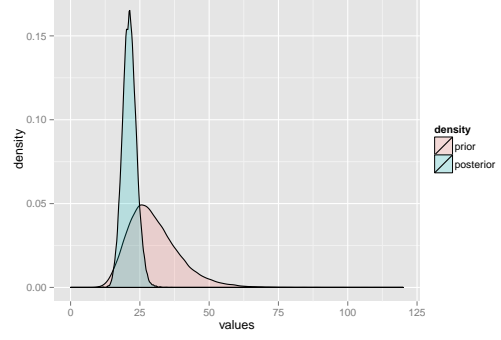
4.4 Spatial predictions

The following figures show the spatial predictions for the spatially varying model parameters, based on the methods from Section 3.4. Every figure in this section on the left panels are based on the uncorrected data set, and the figures on the right hand side are based on the corrected data set.

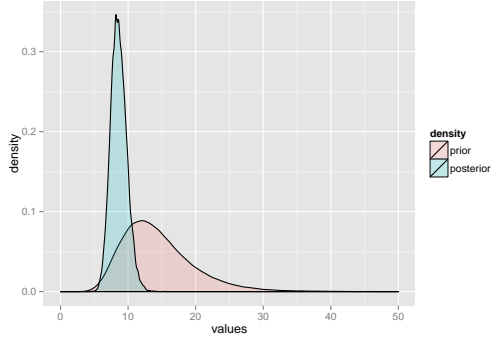
The first row of Figure 8 shows the spatial predictions of the posterior mean of the spatial effect \mathbf{u}_μ on the regular grid \mathcal{G} , based on (3.4). The first row of Figure 8 shows where the spatial model lowers and raises the prediction surface, meaning that the meteorological covariate overestimates and underestimates the extreme precipitation, respectively. The prediction surface is lowered the most in the south-western part of Iceland, but is raised in the south-eastern part. The spatial model yields high positive values for Kvísker, which is known to have the highest observed precipitation in Iceland;



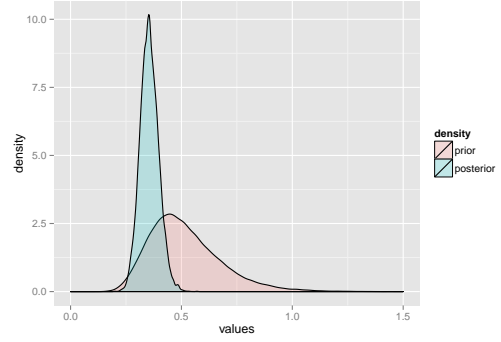
(a) Range of u_μ



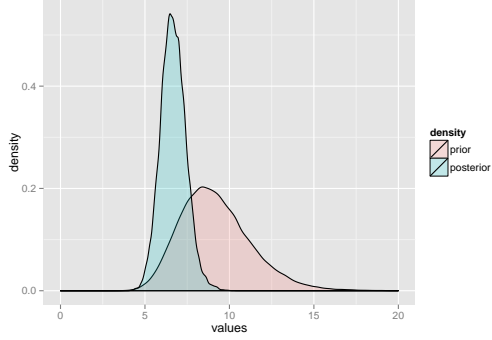
(b) Range of u_τ



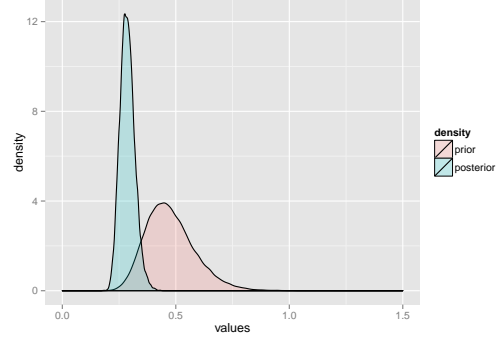
(c) Standard deviation of u_μ



(d) Standard deviation of u_τ



(e) Standard deviation of v_μ



(f) Standard deviation of v_τ

Figure 6: Prior distribution for all the hyperparameters and corresponding posterior distributions based on the uncorrected data set.

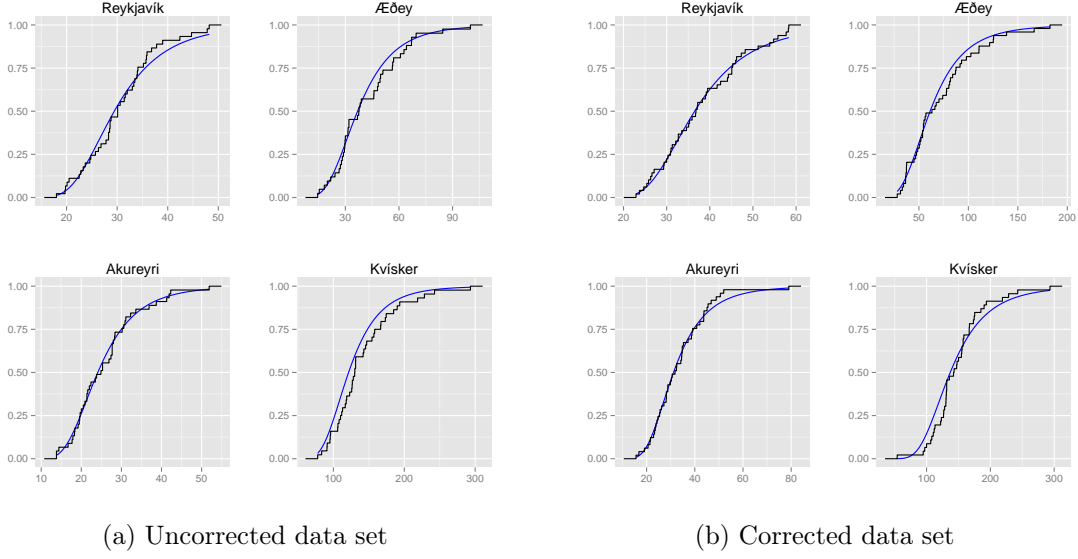


Figure 7: Empirical (black) and model cumulative (blue) distribution functions.

negative value in the south-eastern part close to Reykjavík and values close to zero in the interior of Iceland where there are no observational sites.

The second row of Figure 8 shows the spatial prediction for the standard deviation of the spatial effect \mathbf{u}_μ on the regular grid \mathcal{G} . As expected, the standard deviation is less where there is more data forming sinks in the standard deviation near the observation points. The standard deviation increases at points further away from the observational sites. An interesting artifact of the SPDE modeling also appears. The estimates for the standard deviation are lower for points inside the triangles than in their corresponding edges or vertices in the areas where there are no data. This happens due to the linear basis function and because the standard deviation is estimated based on (3.4).

The third row of Figure 8 shows the spatial prediction for the posterior mean of the location parameter $\boldsymbol{\mu}_\mathcal{G}$, based on (3.5), on the regular grid \mathcal{G} . The figure shows that the mean of the extreme precipitation is at its highest in the south-eastern parts of Iceland, in the vicinity of the southern side of Vatnajökull glacier. This is to be expected, as the spatial gradient at the southern side of Vatnajökull increases rapidly moving from the nearby coastline to the top of the glacier. Due to these topographical properties and the fact that humid air blows in from the southern shoreline towards the roots of the glacier, the physical law of orographic precipitation predicts high precipitation. However, the areas north of Vatnajökull and in the middle of the country are known to be in a rain shadow by the same meteorological law. That is again in line with the results seen in Figure 8.

Figure 9 shows the spatial prediction for the scale parameter on a logarithmic scale, and is arranged in the same manner as Figure 8. In the first row, we see that the spatial model raises the spatial surface in the south-eastern part and lowers it in the south-

western part. In the second row, similar results appear for the standard deviation as for the location parameter. In the third row, we can see that the estimates for the scale parameter are highest along the south-eastern coastline.

Spatial predictions for the posterior mean of the 0.95 quantile of the generalized extreme value distribution can be seen in Figure 10. The predictions are based on the methods discussed in Section 3.6 and can be interpreted as the 20-year precipitation event. The results reflect the previous posterior results about the location and scale parameters. For example, the highest predicted 20-year precipitation events are along the south and south-eastern coastlines, while the lowest predicted events are in the interior of Iceland. The maximum of the predicted values is at the southern side of Hvannadalshnjúkur and is approximately 400 mm per 24 hours, based on the uncorrected data set. The peak of Hvannadalshnjúkur, which is a part Vatnajökull, is the highest point of Iceland and is approximately 10 km north west of the observational site Kvísker. These results demonstrate the effects of orographic precipitation as Kvísker is around 30 m above sea level and the highest point of Hvannadalshnjúkur is roughly 2110 m above sea level. Furthermore, the 20-year precipitation event in Reykjavík and surroundings is predicted close to 54 mm per 24 hours and in the vicinity of Akureyri it is around 42 mm, based on the uncorrected data set.

5 Discussion

The statistical methodology presented in this paper is in line with 'going beyond mean regression' discussed by Kneib [2013]. Working within the LGM framework yields a flexible and computationally efficient way of doing statistical modeling, rather than only doing mean regression. Furthermore, the framework also allows scientific knowledge to be assimilated into a statistical model in a natural way, as presented above.

The presented modeling framework is modular by design. This means that other likelihoods can be chosen for the observations without the need to change the structure of the spatial model at the latent level. A different likelihood choice for extreme precipitation is, for example, the peak over threshold methods with the Generalized Pareto Distribution as presented in [Cooley et al., 2007]. The advantages of that approach is that more of the data can be incorporated into the modeling, depending on the choice of threshold. However, this approach was not chosen for the modeling in this paper, as choosing a common threshold for each station is somewhat unrealistic for the data sets explored in this as this paper. For example, in Figure 2, it can be seen that highest observed value in Reykjavík is lower than the lowest observed value in Kvísker.

Another likelihoods choices are copula likelihoods with marginals based on the generalized extreme value distribution. Likelihoods constructed with the t -copula have been explored for example by Davison et al. [2012] to model extreme precipitation and present an appealing choice to model the probabilistic dependence of the observations. The approach is well suited for simulating spatial realizations of the extreme precipitation for the next year or unobserved sites for an observed year. Although the copula likelihood can be implemented within the presented modeling framework, there were two main rea-

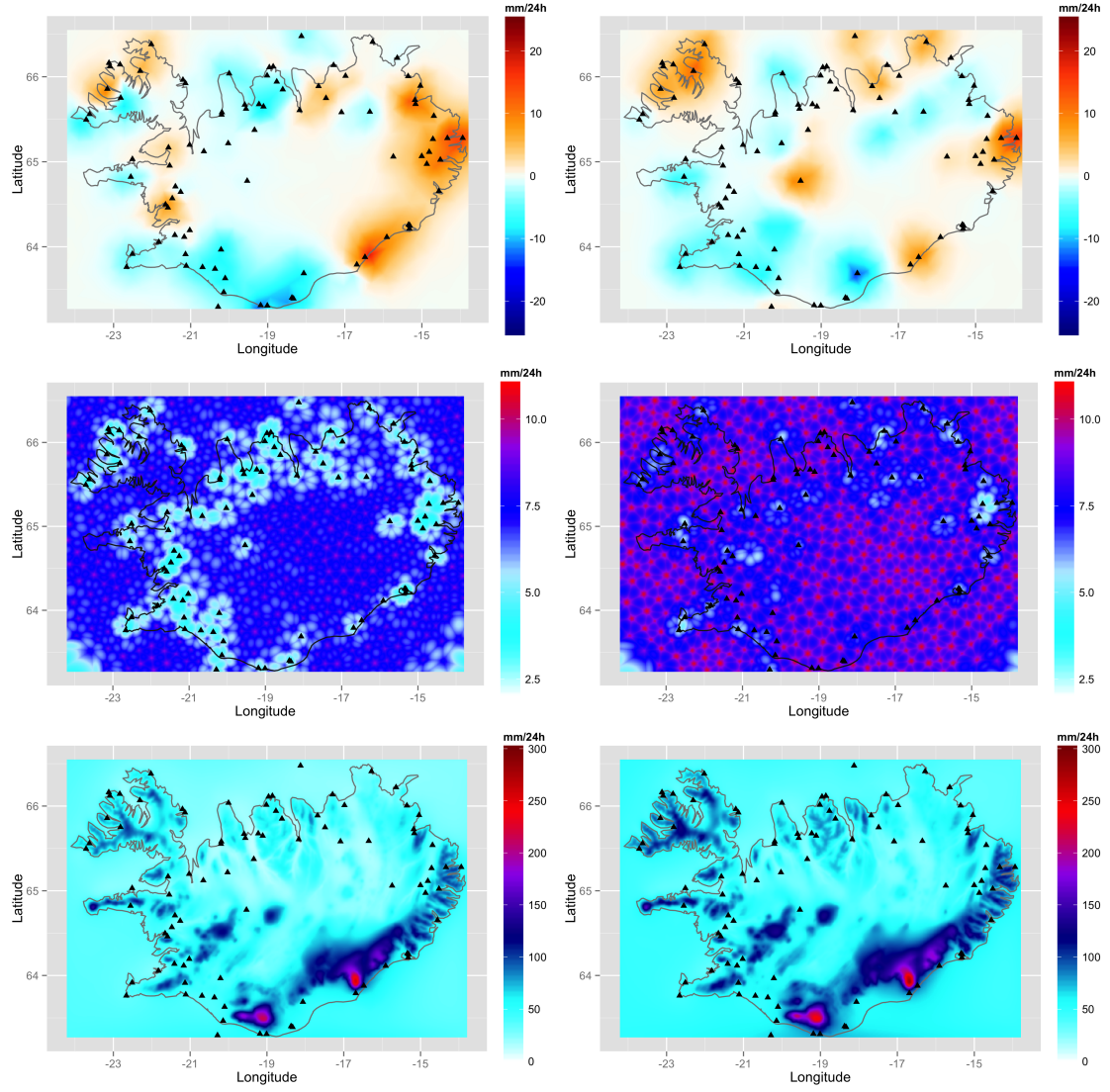


Figure 8: The first, second and third rows show posterior mean of \mathbf{u}_μ , posterior standard deviation of \mathbf{u}_μ and posterior mean of the location parameter μ_G on the regular grid \mathcal{G} , respectively. The left and right panels are based on the uncorrected and corrected data sets, respectively.

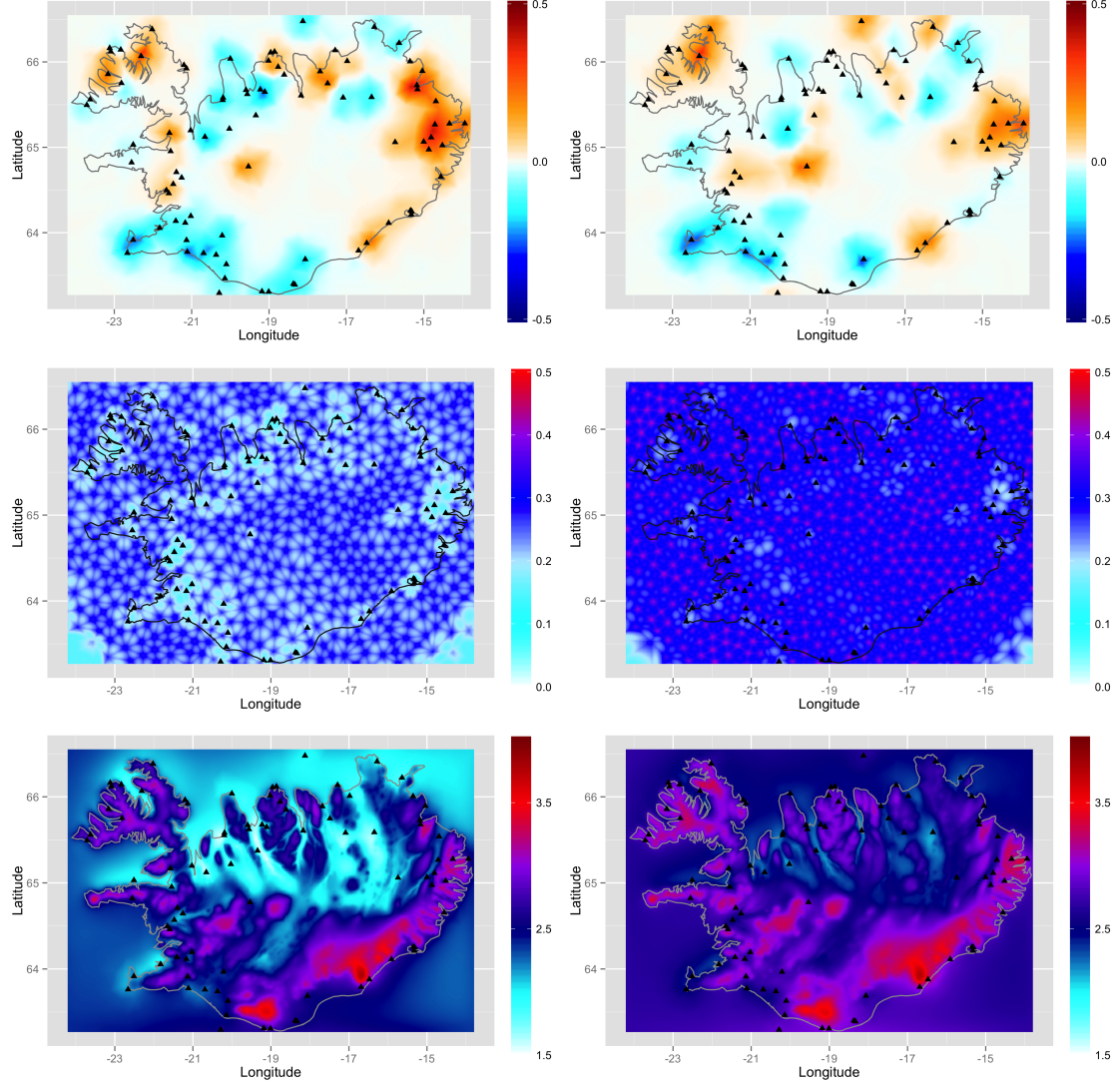


Figure 9: The first, second and third rows show posterior mean of \mathbf{u}_τ , posterior standard deviation of \mathbf{u}_τ and posterior mean of the location parameter τ_G on the regular grid \mathcal{G} , respectively. The left and right panels are based on the uncorrected and corrected data sets, respectively.

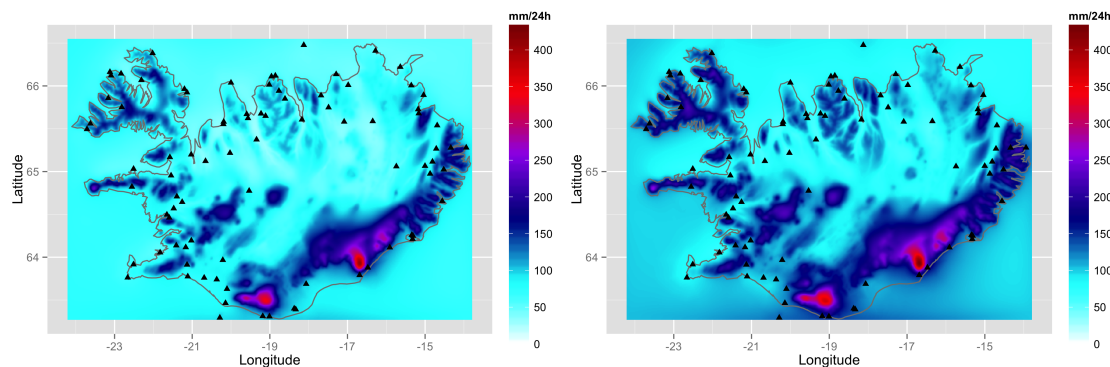


Figure 10: The posterior mean of the 0.95 quantile of the generalized extreme value distribution. The left and right panels are based on the uncorrected and corrected data sets, respectively.

sons why it was not chosen in this paper. First, the main purpose is to give spatial predictions of marginal quantiles using the SPDE approach but not to simulate spatial realizations of the extremal surfaces. Secondly, many of the observational sites had missing observations, which presents further computational difficulties with the copula based likelihood.

At the latent level of the model, we have implemented a stationary SPDE spatial models. However, the results indicated a non-stationary behavior in some of the mountainous regions, for example near Kvísker, as discussed in Section 4.2. Non-stationary SPDE spatial models have been proposed [Fuglstad et al., 2013, Ingebrigtsen et al., 2013], and yield an appealing modeling option for non-stationary spatial fields. However, due to the sparsity of the observational sites, these models were not implemented.

Acknowledgements

The authors would like to thank the University of Iceland Doctoral Fund, University of Iceland Research Fund and Landsvirkjun Energy Research Fund which supported the research. The authors would also like to thank the Icelandic Meteorological Office, in particular, Dr. Philippe Crochet for providing the data and valuable discussions. The authors give their thanks to the Nordic Network on Statistical Approaches to Regional Climate Models for Adaptation (SARMA), especially Prof. Peter Guttorp, for providing travel support. Furthermore, the authors give their thanks to the Department of Mathematical Sciences at the Norwegian University of Science and Technology for hosting Óli Páll Geirsson several times, and special gratitude to Prof. Håvard Rue for his invitation and valuable conversations.

References

- RE Benestad, D Nychka, and LO Mearns. Spatially and temporally consistent prediction of heavy precipitation from mean values. *Nature Climate Change*, 2(7):544–547, 2012.
- Stephen P Brooks and Andrew Gelman. General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics*, 7(4):434–455, 1998.
- Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.
- Daniel Cooley, Douglas Nychka, and Philippe Naveau. Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479):824–840, 2007.
- Philippe Crochet, Tómas Jóhannesson, Trausti Jónsson, Oddur Sigurðsson, Helgi Björnsson, Finnur Pálsson, and Idar Barstad. Estimating the spatial distribution of precipitation in iceland using a linear model of orographic precipitation. *Journal of Hydrometeorology*, 8(6):1285–1306, 2007.
- Anthony C Davison, SA Padoan, and Mathieu Ribatet. Statistical modeling of spatial extremes. *Statistical Science*, 27(2):161–186, 2012.
- Pierre Delfiner et al. *Geostatistics: modeling spatial uncertainty*, volume 497. Wiley.com, 2009.
- Peter J Diggle, JA Tawn, and RA Moyeed. Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(3):299–350, 1998.
- EJ FØrland and Norske Meteorologiske Institutt. *Manual for operational correction of Nordic precipitation data*. Norwegian Meteorological Institute, 1996.
- Geir-Arne Fuglstad, Daniel Simpson, Finn Lindgren, and Håvard Rue. Non-stationary spatial modelling with applications to spatial prediction of precipitation. *arXiv preprint arXiv:1306.0408*, 2013.
- O.P. Geirsson, D. Simpson, and Hrafnkelsson. An mcmc split sampler for latent gaussian models. Technical report, University of Iceland, 2014.
- Peter Guttorp and Tilmann Gneiting. Studies in the history of probability and statistics xlix on the matern correlation family. *Biometrika*, 93(4):989–995, 2006.
- Birgir Hrafnkelsson, Jeffrey S Morris, and Veerabhadran Baladandayuthapani. Spatial modeling of annual minimum and maximum temperatures in iceland. *Meteorology and Atmospheric Physics*, 116(1-2):43–61, 2012.

- Rikke Ingebrigtsen, Finn Lindgren, and Ingelin Steinsland. Spatial models with explanatory variables in the dependence structure. *Spatial Statistics*, (0):–, 2013. ISSN 2211-6753. doi: <http://dx.doi.org/10.1016/j.spasta.2013.06.002>. URL <http://www.sciencedirect.com/science/article/pii/S2211675313000377>.
- Thomas Kneib. Beyond mean regression. *Statistical Modelling*, 13(4):275–303, 2013.
- Leoanhard Knorr-Held and Håvard Rue. On block updating in markov random field models for disease mapping. *Scandinavian Journal of Statistics*, 29(4):597–614, 2002.
- Finn Lindgren, Håvard Rue, and Johan Lindström. An explicit link between gaussian fields and gaussian markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011.
- Håvard Rue. Fast sampling of gaussian markov random fields. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):325–338, 2001.
- Havard Rue and Leonhard Held. *Gaussian Markov random fields: theory and applications*. CRC Press, 2005.
- Håvard Rue, Sara Martino, and Nicolas Chopin. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, 71(2):319–392, 2009.
- Huiyan Sang and Alan E Gelfand. Hierarchical modeling for extreme values observed over space and time. *Environmental and Ecological Statistics*, 16(3):407–426, 2009.
- Ronald B Smith and Idar Barstad. A linear theory of orographic precipitation. *Journal of the Atmospheric Sciences*, 61(12), 2004.
- Claudia Tebaldi and Bruno Sansó. Joint projections of temperature and precipitation change from multiple climate models: a hierarchical bayesian approach. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1):83–106, 2009.
- Sakari M Uppala, PW Kållberg, AJ Simmons, U Andrae, V Bechtold, M Fiorino, JK Gibson, J Haseler, A Hernandez, GA Kelly, et al. The era-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612):2961–3012, 2005.

6 Appendix

6.1 Convergence diagnostic plots

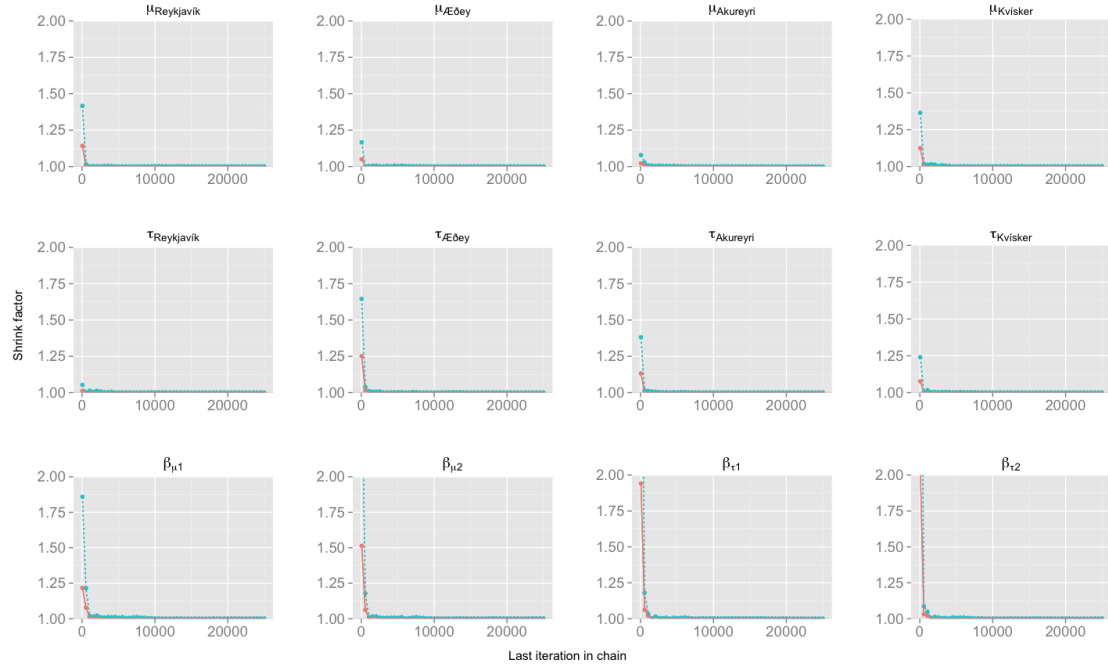


Figure 11: The first two rows show Gelman–Rubin statistics as a function of iterations after burn-in, for μ and τ , respectively, at Reykjavík, Æðey, Akureyri and Kvísker, based on the uncorrected data set. The third row shows a Gelman–Rubin plot for the coefficients of the covariates.

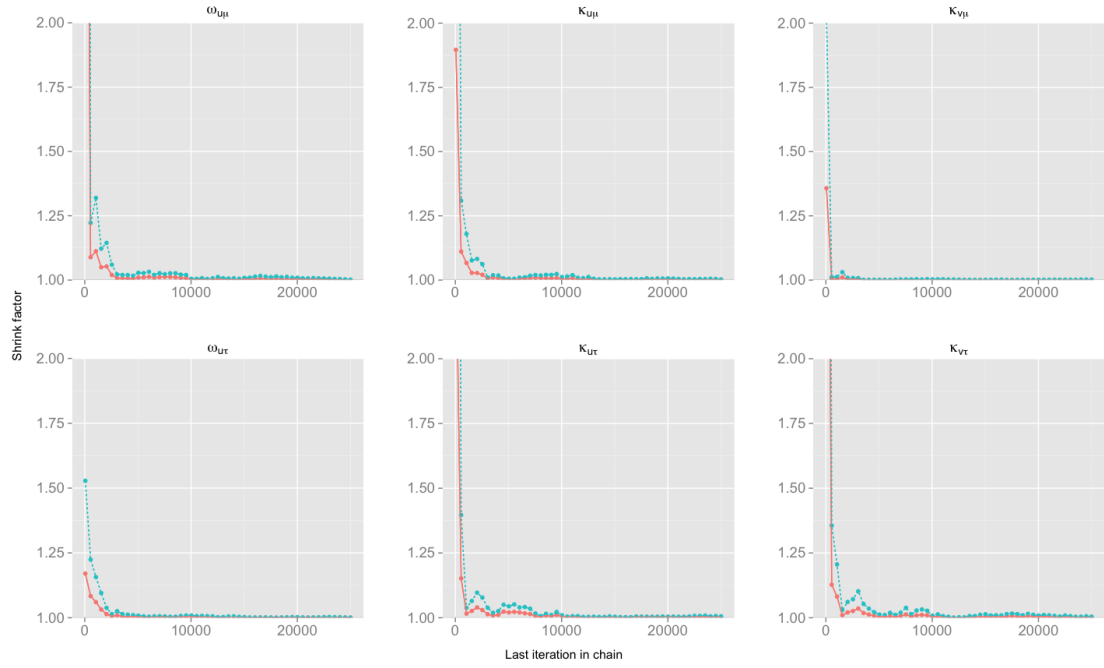


Figure 12: The figure shows Gelman–Rubin statistics as a function of iterations after burn-in, for all the hyper parameters $\kappa_{u\mu}$, $\omega_{u\mu}$, $\kappa_{u\tau}$, $\omega_{u\tau}$, $\kappa_{v\mu}$ and $\kappa_{v\tau}$ based on the uncorrected data set

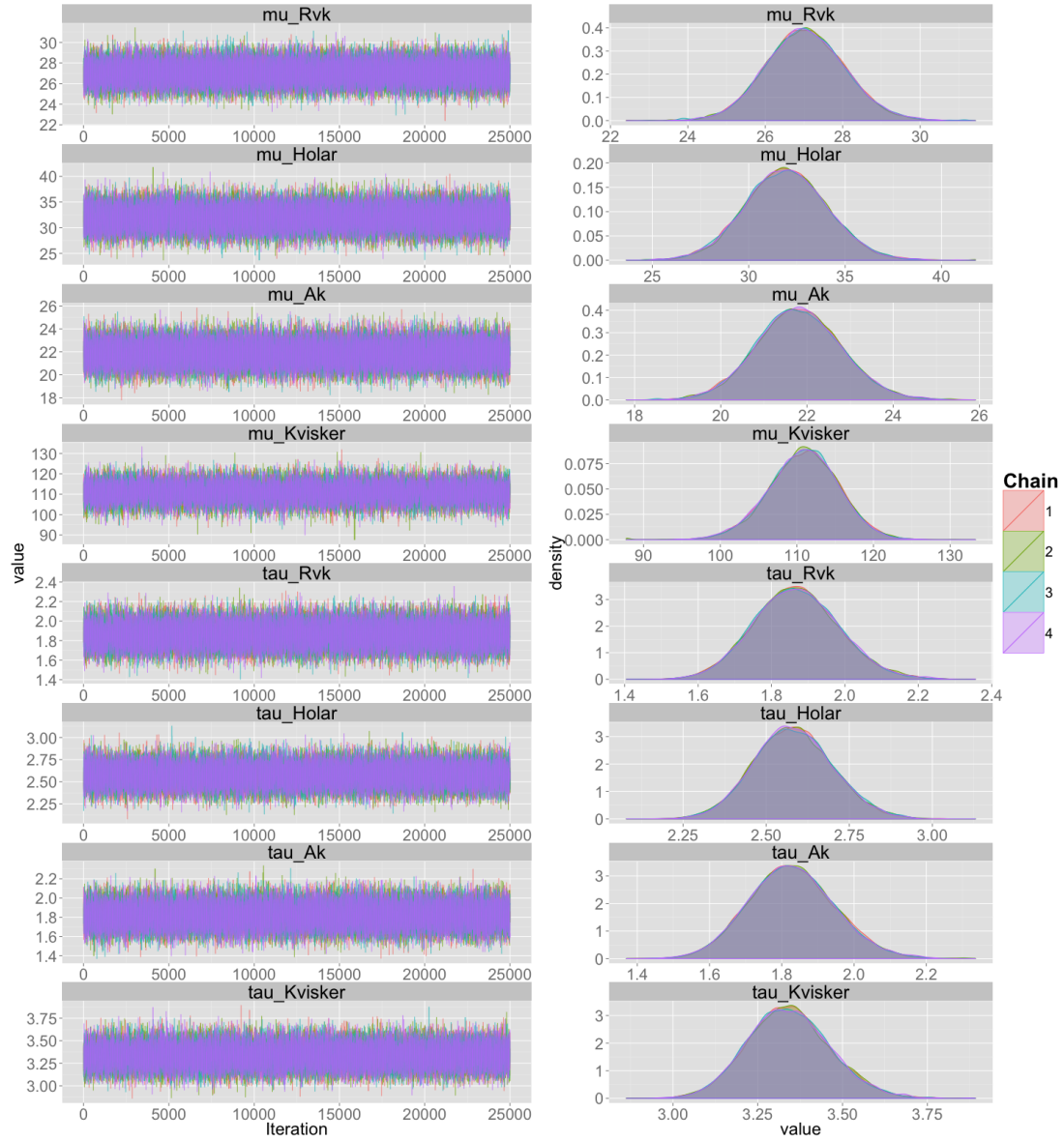


Figure 13: Posterior trace and density plots of μ and τ at Reykjavík, Aedey, Akureyri and Kvísker, based on the uncorrected data set

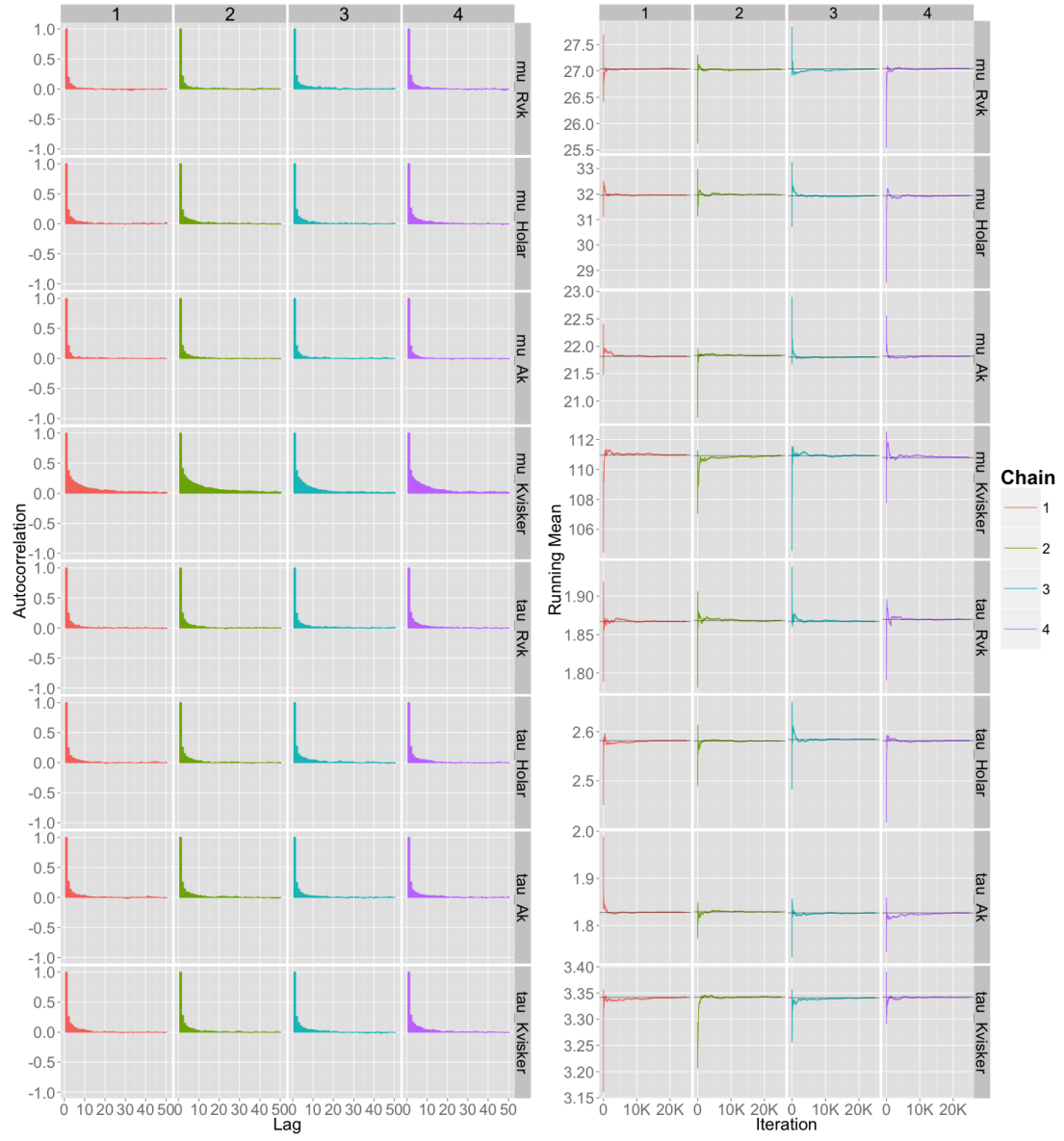


Figure 14: Auto correlation and running mean plots of μ and τ at Reykjavík, Æðey, Akureyri and Kvísker, based on the uncorrected data set

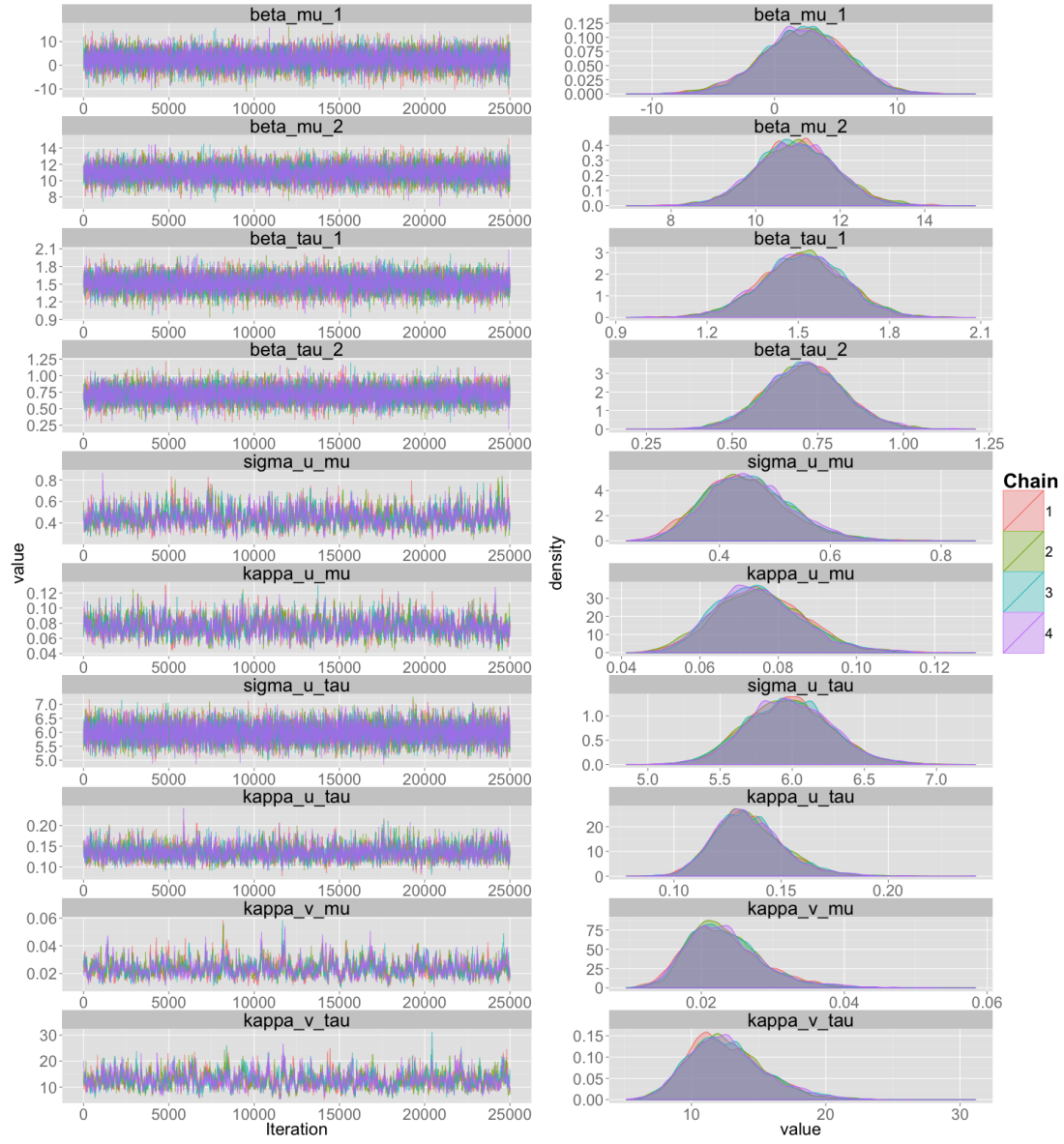


Figure 15: Posterior trace and density plots of the coefficients of the covariates β_μ and β_τ ; and of the all hyperparameters, based on the uncorrected data set.

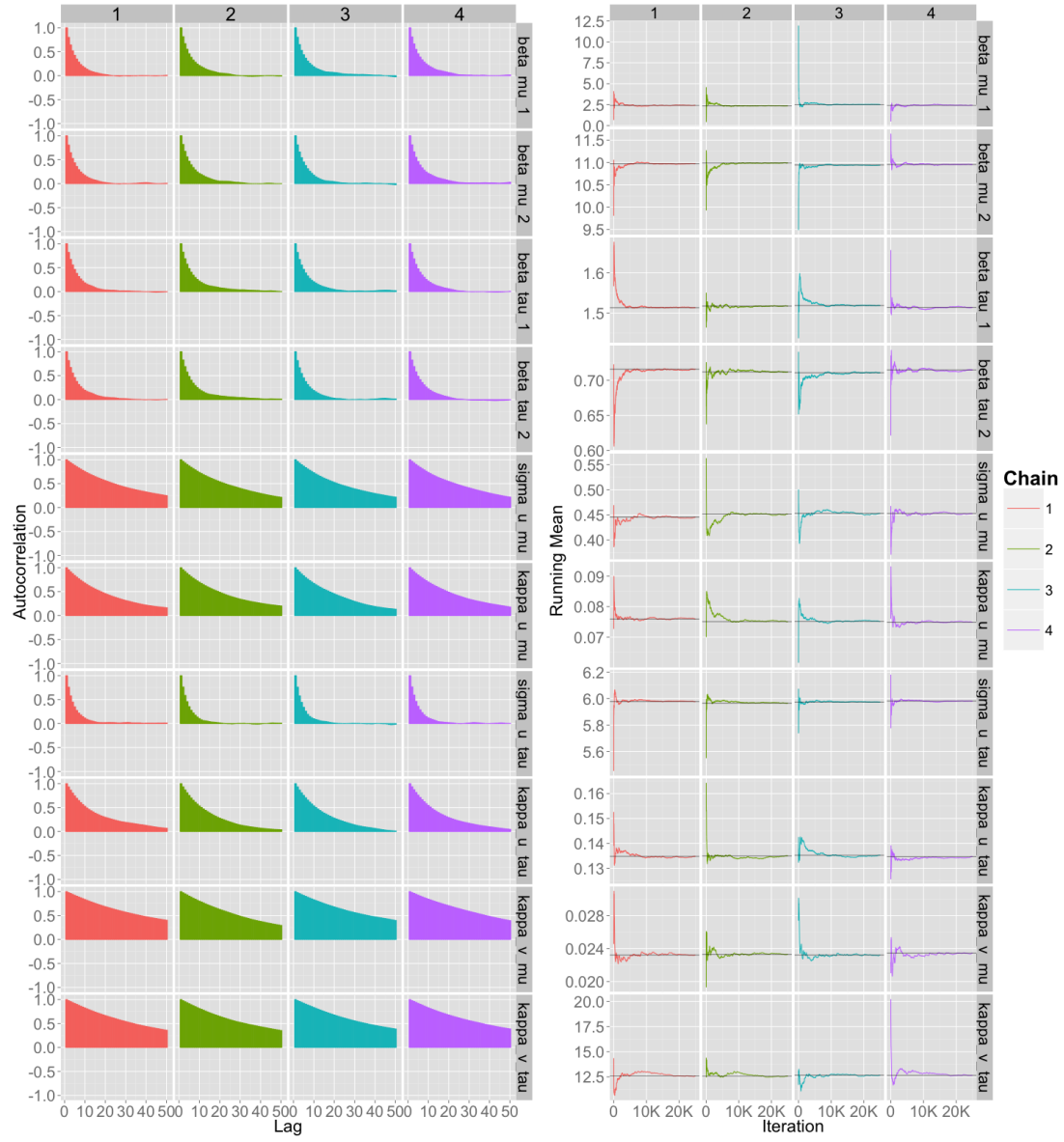


Figure 16: Auto correlation and running mean plots of the coefficients of the covariates β_{μ} and β_{τ} ; and of the all hyperparameters, based on the uncorrected data set.